آمار و احتمال مهندسی
فصل اول آمار توصـــیفی

مقدمه

امروزه در علوم پایه و مهندسی به دفعات نـاز به جمع آوری اطلاعات در مورد مجموعه‌های از اشیاء و یا انسانها داریم که این امر بر عهده علم آمار می‌باشد و با کمک آمار توصیفی می‌توان اطلاعات جمع آوری شده را به صورتی منظم گرد آوری نمود بطوریکه بتوان با یک نگاه اجمالی به نتایج بدست آمده یک دید کلی نسبت به کل داده‌ها بدست آورد.

در این فصل به چگونگی جمع آوری، اطلاعات و تجزیه و تحلیل آن با کمک نمودارها و پارامترهای مرکزی و پراکندگی می‌پردازیم.

۱-۱ تعاریف اولیه

جامعه آماری : به مجموعه‌ای از اشیاء یا افراد که حداقل یک ویژگی مشترک آنها مورد مطالعه قرار می‌گـیرد. جامعه آماری می‌گوییم. ویژگی‌های مشترک یک جامعه آماری از عضوی به عضو دیگر تغییر می‌کند. که آنها را متغیر می‌نامند. متغیرها معمولاً به دو نوع تقسیم می‌شوند:

متغیر کمی : متغیرهایی می‌باشند که معمولاً قابل اندازه‌گیری هستند و می‌توان مقدار آنها را به صورت عددی نمایش داد مثل مقدار وزن، قد، حجم و ....

۱- متغیر کیفی : متغیرهایی می‌باشند که مستقیماً توسط اعداد و ارقام قابل اندازه‌گیری نیستند. مثل گروه خونی، شغل، رنگ چشم و ... که برای اندازه‌گیری این متغیرها به آنها عددی نسبت می‌دهیم.

متغیرهای کمی خود بر دو نوع هستند.

۱- گسسته : متغیرهایی که بین دو مقدار متصور آنها هیچ عدد دیگری وجود نداشته باشد.

۲- پیوسته : متغیرهایی که بین هر دو مقدار متصور آنها همواره عددی دیگری وجود دارد. مثل وزن یا طول و قد افراد.

پس از جمع آوری داده‌ها برای رسیدن به اهداف مورد نیاز به بررسی و تجزیه و تحلیل داده‌ها داریم که برای این منظور ابتدا داده‌ها را در یک جدول تنظیم و طبقه‌بندی می‌کنیم و سپس با استفاده از نمودارهای آماری نحوه توزیع داده‌ها را نمایش می‌دهیم و در نهایت داده‌ها را با کمک چند عدد به نام شاخص یا آماره خلاصه می‌کنیم.

۲-۱ جدول آماری

در جداول آماری علاوه بر داده‌ها، تعداد و درصد تکرار آنها نمایش داده می‌شود بطوریکه با یک نگاه به جدول می‌توان اطلاعات مفیدی در مورد پراکندگی و توزیع داده‌ها بدست آورد. یکی از متداول‌ترین جداول آماری جدول فراوانی است در جداول فراوانی همواره موارد زیر را خواهیم داشت:

۱- فراوانی: با شمردن تعداد دفعات تکرار هر داده در میان تمامی داده‌ها فراوانی آن داده بدست می‌آید. فرض کنید N داده داشته باشیم با قرار دادن داده‌های مشابه در یک دسته، در نهایت K طبقه خواهیم داشت $(K \le N)$ که تعداد داده‌ها در هر دسته را فراوانی آن داده می‌نامیم. فراوانی طبقه i ام را با f_i نمایش می‌دهیم که $1 \le i \le k$ و $1 \le f_i \le N$ و $\sum_{i=1}^{k} f_i = N$ (N تعداد کل داده‌ها)

۲- فراوانی نسبی : عبارتست از حاصل تقسیم فراوانی هر طبقه بر تعداد کل داده‌ها که آنرا با $i = 1, \cdots k$ ، $r_i = \frac{f_i}{N}$ نمایش می‌دهیم.

$$\sum_{i=1}^{k} r_i = \sum_{i=1}^{k} \frac{f_i}{N} = \frac{1}{N} \sum_{i=1}^{k} f_i = 1$$

هم چنین اگر فراوانی نسبی را در عدد ۱۰۰ ضرب کنید درصد وقوع هر داده را در میان کل داده‌ها بدست می‌آورید.

۳- فراوانی تجمعی نسبی : حاصل جمع فراوانی هر طبقه با طبقات قبل از آنرا فراوانی تجمعی می‌نامیم و به $(1 \le i \le k)$ g_i نمایش می‌دهیم.

فراوانی تجمعی طبقه iام $= g_i = f_1 + f_2 + \cdots + f_1 = \sum_{j=1}^{i} f_j$

به همین ترتیب حاصل جمع فراوانی نسبی هر طبقه با طبقات قبل آنرا فراوانی تجمعی نسبی می‌نامیم و به $s_i \ (1 \le i \le k)$ نمایش می‌دهیم

فراوانی تجمعی طبقه iام $= s_i = r_1 + r_2 + \cdots + r_1 = \sum_{j=1}^{i} s_j = s_i = \frac{1}{N} g_i$

توجه : در محاسبه g_i و s_i اگر x_i نماینده طبقه iام باشد می‌بایستی داده‌ها به صورت مرتب و از کوچک به بزرگ در جدول قرار داده شوند بطوریکه داشته باشیم $x_1 < x_2 < \cdots x_n$.

۱-۲-۱ جدول فراوانی برای داده های گسسته

در مثال زیر چگونگی تشکیل جدول فراوانی برای داده‌های گسسته را بیان می‌کنیم.

مثال ۱ : داده‌های زیر تعداد فرزندان تحت پوشش بیمه در ۳۰ خانواده را نشان می‌دهد.

5 , 4 , 2 , 1 , 1 , 1 , 3 , 3 , 3 , 2 , 0 , 5 , 0 , 2 , 2 , 2 ,2 , 4 , 4 , 1 , 0 , 1 , 1 , 3 , 2 , 5 , 2 , 1 , 0 , 3

ابتدا داده‌ها را از کوچک به بزرگ برای تشکیل جداول فراوانی مرتب می‌کنیم:

0 , 0 , 0 ,0 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 2 , 2 , 2 , 2 , 2 ,2 , 2 , 2 , 3 , 3 , 3 , 3 , 4 , 4 , 4 , 4 , 5 , 5 , 5

در این صورت ۶ طبقه بدست خواهد آمد. نماینده هر طبقه در ستون x_i و فراوانی ان در ستون f_i نوشته می‌شود به این ترتیب جدول زیر را خواهیم داشت:

x_i	f_i	r_i	g_i	s_i
۰	۴	۰/۱۳	۴	۰/۱۳
۱	۷	۰/۲۳	۱۱	۰/۳۶
۲	۸	۰/۲۶	۱۹	۰/۶۲
۳	۴	۰/۱۳	۲۳	۰/۷۵
۴	۴	۰/۱۳	۲۷	۰/۸۸
۵	۳	۰/۱	۳۰	۱
جمع	۳۰	۱		

با توجه به جدول می‌توان به نتایج زیر رسید:

۱- با توجه به عدد ۰/۲۶ در ستون فراوانی نسبی می‌توان نتیجه گرفت که ۲۶ درصد از خانوارها دارای ۲ فرزند می‌باشند.

۲- با توجه به عدد ۰/۷۵ در ستون فراوانی تجمعی نسبی می‌توان نتیجه گرفت که ۷۵ درصد خانوارها حداکثر دارای ۳ فرزند می‌باشند.

۲-۲-۱ جدول فراوانی برای داده‌ها پیوسته

مثال ۲ : داده‌های زیر طول عمر ۴۰ عدد لامپ را نشان می‌دهند که به نزدیک‌ترین عدد صحیح گرد شده‌اند.

11	9	12	15	20	13	14	17	23	22
8	16	17	21	11	18	21	12	11	10
14	13	19	16	15	17	20	8	7	13
15	17	16	14	22	12	11	9	18	19

زمانی که با داده‌های پیوسته کار می‌کنیم داده‌ها را به رده‌های (فاصله‌ها) با طول مساوی تقسیم می‌کنیم و فراوانی داده‌ها را در هر رده بدست می‌آوریم. در این حالت نیاز به ثبت تمامی داده‌ها در جدول نمی‌باشد بلکه هر رده را به صورت مرتب از کوچک به بزرگ در جدول ثبت می‌کنیم.

در این حالت روند به این صورت است که :

۱- از آنجا که هر عدد به نزدیکترین عدد صحیح گرد شده است پس مثلاً عدد ۱۲ در واقع عددی بین (۱۲/۵ و ۱۱/۵) بوده است برای اینکه این مقدار نیز در عملیات وارد شود از مفهوم میزان تغیر پذیری داده‌ها که با s نمایش می‌دهیم استفاده می‌کنیم.

$$S=\frac{\text{واحد گرد شده داده‌ها}}{2}=\frac{1}{2}=0/5$$

۲- دامنه واقعی داده‌ها و کوچکترین و بزرگترین داده را به این ترتیب محاسبه می‌کنیم.

$$\min=\text{کوچکترین داده}-S=7-0/5=6/5$$

$$\max=\text{بزرگترین داده}+S=23+0/5=23/5$$

$$\text{تغیرات داده‌ها}\quad\text{دامنه}\quad R=\max-\min=23/5-6/5=11$$

۳- با توجه به دامنه داده‌ها می‌توان طول هر رده را معین نمود. برای این می‌بایستی دامنه را بر تعداد رده‌ها (که به صورت دلخواه قابل انتخاب است) تقسیم نمود.

$$\text{طول رده}=\omega=\frac{R}{K}\qquad K=\text{تعداد رده‌ها}$$

معمولاً تعداد رده‌ها طوری انتخاب می‌شوند که هر رده حداقل پنج داده را در برداشته باشد می‌توان از فرمول $K=1+3/322\log_{10}^{N}$ برای تعیین تعداد رده استفاده نمود بنابر این:

$$K=1+3/3222\log_{10}^{40}=6/322\cong 7\quad\Rightarrow\quad\omega=\frac{17}{7}=2/4\cong 3$$

۴- حالا کافیست ۷ رده به طول ۳ در ستون اول جدول فراوانی ایجاد کنیم و مجدداً مطابق مثال قبل جدول را کامل کنیم. اولین رده از کوچکترین عنصر آغاز می‌شود.

رده‌ها	خط و نشان	x_i	fi	r_i	g_i	s_i
$6/5-9/5$	𝍸	۸	۵	$0/125$	۵	$0/125$
$9/5-12/5$	𝍸 |||	۱۱	۸	$0/2$	۱۳	$0/325$
$12/5-15/5$	𝍸 ||||	۱۴	۹	$0/225$	۲۲	$0/55$
$15/5-18/5$	𝍸 ||||	۱۷	۹	$0/225$	۳۱	$0/775$
$18/5-21/5$	𝍸		۲۰	۶	$0/15$	۳۷	$0/925$
$21/5-24/5$	|||	۲۳	۳	$0/075$	۴۰	$1/0$
جمع			۴۰			

توجه شماره ۱ :

- x_i معرف نماینده هر رده است که عضو میانی رده می‌باشد.
- چون آخرین رده شامل هیچ عضوی نبود آنرا حذف می‌کنیم مثلاً در این مثال رده ۲۴/۵-۲۷/۵ شامل هیچ عضوی نیست بنابر این حذف می‌شود.

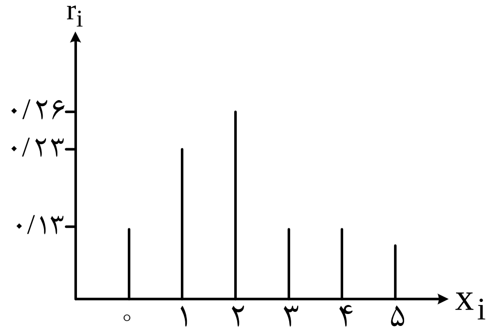
۱-۳ نمودارهای آماری

یکی دیگر از روشهای مناسب برای خلاصه نمودن داده‌های آماری استفاده از نمودار می‌باشد. نمودارها بهترین ابزار برای نمایش نحوه توزیع داده‌ها می‌باشند که در ذیل معروفترین آنها را معرفی می‌کنیم.

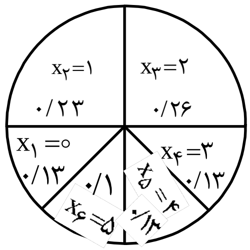
۱-۳-۱ نمودارهای آماری برای داده‌های گسسته

۱- نمودارهای میله‌ای:

در این نمودار محور x ها نمایش دهنده مقادیر داده‌ها و محور y ها نمایش دهنده فراوانی نسبی می‌باشد. چون فراوانی نسبی داده‌های هر جامعه آماری مقادیر بین صفر و یک را می‌گیرد بنابر این می‌توان چندین نمودار را با یکدیگر مقایسه نمود. شکل نمودار میله‌ای را برای مثال ۱ نمایش می‌دهد.



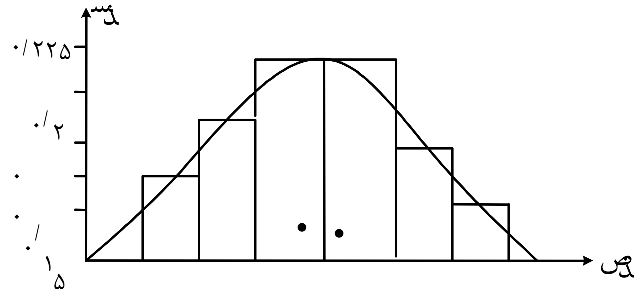
۳- نمودار دایره‌ای : این نمودار درون دایره‌ای رسم می‌شود که به قطاع‌هایی تقسیم شده است و هر قطاع متناسب با مساحت خود معرف یکی از فراوانی‌های نسبی در جدول فراوانی می‌باشد. شکل زیر نمایش دهنده نمودار دایره‌ای برای مثال ۱ می‌باشد.



۱-۳-۲ نمودارهای آماری برای داده‌های پیوسته

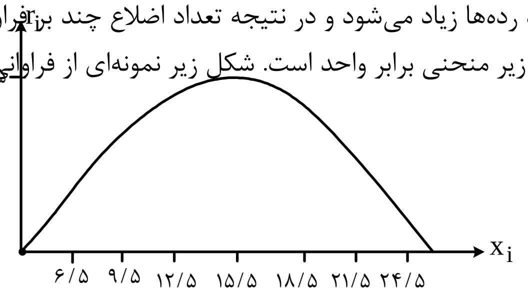
همانطور که در جدول فراوانی مشاهده کردید تفوت عمده رده‌های پیوسته با گسسته در بکارگیری رده‌ها می‌باشد این امر در نمودارها نیز کاملاً صادق است. در این قسمت سه نمودار را که برای نمایش داده‌های پیوسته بکار می‌روند معرفی می‌کنیم.

۱- هیستوگرام (نمودار ستونی) : این نمودار کاملاً مشابه نمودار میله‌ای می‌باشد با این تفاوت که محور xها نمایش دهنده رده‌های جدول فراوانی است و به ازای هر رده مستطیلی که عرض آن یک واحد و ارتفاع آن معادل فراوانی نسبی آن رده است رسم می‌شود. بنابر این مجموع مساحتهای مستطیل‌های رسم شده برابر واحد می‌باشد که همان مجموع کل فراوانی نسبی‌ها است. شکل زیر نمودار هستیوگرام برای مثال ۲ می‌باشد.



۲- چند برابر فراوانی: اگر در نمودار هستیوگرام وسط قاعده‌های بالای مستطیل‌ها را به یکدیگر توسط خطوطی به صورت متوالی متصل کنیم نمودار چند برابر فراوانی بدست می‌آید که در شکل بالا نیز نشان داده شده است. ابتدای خطوط به وسط رده ماقبل و انتهای خطوط به وسط رده مابعد مستطیلهای هستیوگرام متصل می‌شوند به این ترتیب مساحت زیر نمودار چند بر فراوانی مجدداً برابر واحد خواهد بود.

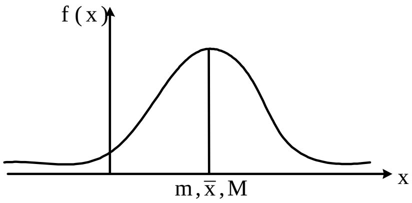
۳- منحنی فراوانی: در صورتی که تعداد داده‌ها زیاد باشد و طول رده‌ها کوچک، تعداد رده‌ها زیاد می‌شود و در نتیجه تعداد اضلاع چند بر فراوانی افزایش یافته و در نهایت مشابه یک منحنی خواهد شد که در این حالت نیز مساحت زیر منحنی برابر واحد است. شکل زیر نمونه‌ای از فراوانی است.

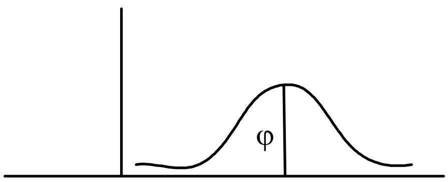


متغیر تصادفی نرمال: اگر مجموعه داده‌های X دارای میانگین $\bar{x}$ و واریانس σ^2 باشند و نمودار آنها از تابع

$$(-\infty < x < +\infty) \qquad f_{(x)} = \frac{1}{\sqrt{2\pi\sigma}}\, e^{\frac{-(x-\bar{x})^2}{2^2}}$$

تبعیت کند می‌گوییم متغیر تصادفی X نرمال می‌باشد. نمودار متغیر تصادفی نرمال نسبت به خط $x=\bar{x}$ متقارن می‌باشد. به شکل زیر توجه کنید.





همچنین با افزایش مقدار δ نمودار در راستای محور xها کشیده تر می‌شود.
با افزایش مقدار $\bar{x}$ نمودار به سمت راست و با کاهش آن به سمت چپ جابجا می‌شود.

۴-۱ پارامترهای مرکزی و پراکندگی

از آنجا که با مطالعه یک جامعه آماری تعداد زیادی داده بدست می‌آوریم و مطالعه روی تک تک یا قسمتی از این داده‌ها مشکل و حتی غیر ممکن است همواره علاقه داریم این داده‌ها را با کمک شاخص‌ها و پارامترهایی خلاصه کنیم تا بتوان با یک نگاه اجمالی به آن یک دید کلی نسبت به کل داده‌ها و جامعه آماری بدست آورد. بنابر این پارامترهای مرکزی و پراکندگی را معرفی می‌کنیم.

۱-۴-۱ پارامترهای مرکزی

عموماً داده‌ها در جامعه آماری یکنوع تجمع .و فشردگی حول یک مقدار خاص از صفت مورد مطالعه را بوجود می‌آورند که این مقدار خاص به عنوان یک پارامتر مرکزی معرفی می‌شود. مهمترین پارامترهای مرکزی عبارتند از: میانگین، میانه و مد (نما)

۱- میانگین: میانگین بر چند نوع است که معروفترین آنها میانگین حسابی، خندسی، همساز (هارمونیک) و درجه دوم می‌باشد.

الف) میانگین حسابی: اگر داده های $x_1, x_2, \ldots, x_k$ به ترتیب دارای فراوانی $f_1, f_2, \ldots, f_k$ باشند و تعداد کل این داده‌ها N باشد $\bar{X}$ یا میانگین حسابی به صورت زیر تعریف می‌شود:

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{K} f_i\, x_i$$

برای داده‌های پیوسته x_i را نماینده هر رده در نظر می‌گیریم. به عنوان مثال میانگین حسابی برای مثال ۱ و ۲ عبارتست از:

مثال 1 $: \bar{X} = \frac{1}{30}\,(4\times0+7\times1+8\times2+4\times3+4\times4+3\times5) = 2/2$

مثال 2 $: \bar{X} = \frac{1}{40}\,(8\times5+11\times8+14\times9+17\times9+20\times6+23\times3) = 14/9$

ب) میانگین هندسی: در صورتی که همگی داده‌ها مثبت باشند میانگین هندسی به صورت زیر محاسبه می‌شود.

$$G=\left(\prod_{i=1}^{N} x_i^{f_i}\right)^{\frac{1}{N}}$$

ج: میانگین همساز: (همنوا یا هارمونیک) : اگر هیچکدام از داده ها صفر نباشد می‌توان میانگین همساز را از طریق فرمول محاسبه نمود.

$$H=\frac{N}{\sum_{i=1}^{n}\frac{f_i}{x_i}}$$

در حالت کلی رابطه $H \le G \le \overline{X}$ بین این سه میانگین برقرار است. که حالت تساوی زمانی رخ می دهد که همگی داده‌ها برابر باشند.

۲- میانه: با مرتب نمودن داده‌ها به صورت غیر نزولی عدد m را بعنوان میانه آنها در نظر می‌گیریم اگر تقرایباً نیمی از داده‌ها سمت چپ آن و نیمی دیگر در سمت راست آن قرار داشته باشند.

همچنین با توجه به جدول فراوانی، میانه کوچکترین مقدار x است که فراوانی تجمعی آن بیشتر یا مساوی با $\frac{N}{۲}$ باشد.

- برای داده‌های گسسته در صورتی که داده‌ها را به صورت غیر نزولی مرتب کنیم اگر تعداد داده‌ها فرد باشد میانه داده وسطی یا به عبارتی $m=x_{\left(\frac{n+1}{2}\right)}$ خواهد بود و اگر تعداد داده‌ها زوج باشد، میانه میانگین دو داده وسطی می‌باشد

$$m=\frac{x_{\left(\frac{n}{2}\right)}+x_{\left(\frac{n}{2}+1\right)}}{2} \qquad ; m=\frac{x_{\left(\frac{n}{2}\right)}+x_{\left(\frac{n}{2}+1\right)}}{2}=\frac{2+2}{2}=4$$

- برای داده‌های پیوسته ابتدا می‌بایستی رده میانه دار را پیدا نمود. رده میانه دار اولین رده‌ای است که فراوانی تجمعی نسبی آن از ۰/۵ بیشتر است. میانه عددیست که در این رده قرار دارد. حال برای محاسبه آن از فرمول زیر استفاده می‌کنیم:

$$m=L_{0/5}+\frac{(0/5N-g_{0/5})\omega}{f_{0/5}}$$

که در آن :

$L_{0/5}$: کران پایین رده میانه دار

N : تعداد داده ها

$g_{0/5}$: فراوانی تجمعی رده قبل از رده میانه دار

$f_{0/5}$: فراوانی رده میانه دار

ω : طول رده

به عنوان مثال میانه برای مثال ۱ با توجه به تعداد زوج داده‌ها عبارتست از:

برای مثال ۲ نیز میانه برابر است با:

$$m=12/5+\frac{(0/5\times40-13)3}{9}=14/8$$

۳- مد یا نما: داده‌ای که فراوانی آن از سایر داده‌ها بیشتر باشد مد یا نما گفته می‌شود و با M نمایش داده می‌شود.

الف) روش محاسبه مد برای داده‌های گسسته: پس از بدست آوردن فراوانی داده، داده‌های که فراوانی آن از بقیه بیشتر باشد مد می‌باشد در این صورت سه حالت زیر پیش می‌آید: (که در مثال زیر به آن می‌پردازیم)

مثال ۳ : برای داده‌های مثال ۱ تنها عدد ۲ دارای بیشترین فراوانی می‌باشد. پس عدد ۲ مقدار مد خواهد بود. اما برای داده‌های ۵ و ۴ و ۴ و ۲ و ۲ و ۲ و۱ و ۱ و ۱ چون فراوانی ۱ و ۲ برابر است و این دو عدد به صورت متوالی قرار گرفته‌اند در این حالت مد برابر است با میانگین این دو عدد:

$$M=\frac{1+2}{2}=\frac{3}{2}$$

اما برای داده‌های ۴ و ۴ و۲ و ۲ و ۱ و ۱، فراوانی ۱ و ۴ مساوی می‌باشد ولی مجاور یکدیگر نیستند دو مقدار برای مد خواهیم داشت که عبارتند از:

$$M_1 = 1 \qquad M_2 = 4$$

توجه کنید در صورتی که فراوانی همه داده‌ها برابر باشد داده‌ها بدون مد در نظر گرفته می‌شوند.

ب) محاسبه مد برای داده‌های پیوسته: رده‌ای که فراوانی آن از سایر رده‌ها بیشتر است را به عنوان رده نمایی انتخاب می‌کنیم در این حالت می‌توان نماینده رده را به عنوان مد یا نما انتخاب کرد. اما برای محاسبه دقیق تر از فرمول زیر نیز می‌توان استفاده نمود:

$$M = L_M + \left(\frac{D_1}{D_1 + D_2}\right)\omega$$

L_M: کران پایین رده نمایی.
D_1: اختلاف فراوانی‌های نسبی رده نمایی و رده قبل از آن.
D_2: اختلاف فراوانی‌های نسبی رده نمایی و رده بعد از آن.
ω: طول رده.
در مثال ۲ داریم:

$$M_1 = 14 \quad ; \qquad M_2 = 17$$

$$M = \frac{14+17}{2} = 15/5$$

و یا بصورت دقیق تر:

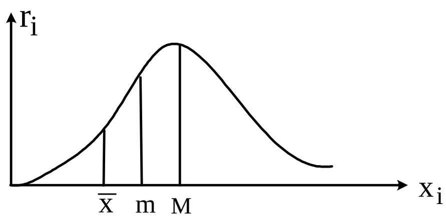
$$M_1 = 12/5 + \left(\frac{0/025}{0/025+0}\right) \times 3 = 15/5$$

$$M_2 = 15/5 + \left(\frac{0}{0+0/075}\right) \times 3 = 15/5$$

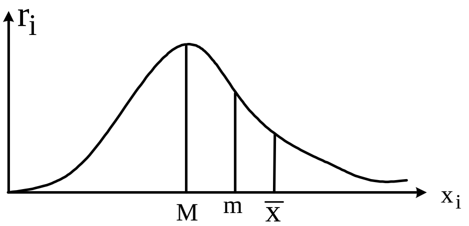
$$\Rightarrow M = 15/5$$

۱-۵ چولگی توزیع فراوانی

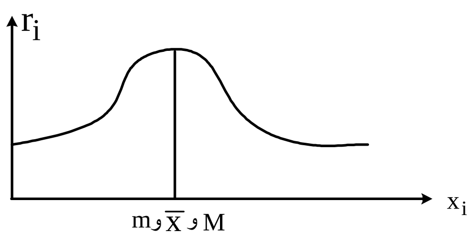
با رسم نمودار فراوانی داده‌های پیوسته عموماً یکی از اشکال زیر بدست می‌آید. که چگونگی قرار گرفتن میانگین و میانه و مد را در اشکال زیر مشاهده می‌شود.



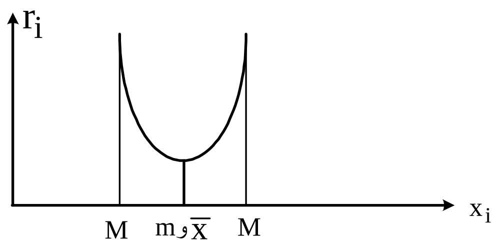
چوله به‌چپ $\overline{X} < m < M$



چوله به‌راست $M < m < \overline{X}$



متقارن $M = \overline{X} = m$



متقارن $m = \overline{X} = M$

برای محاسبه میزان عدم تقارن منحنی، شاخصی به نام B_1 (ضریب چولگی) وجود دارد که در فصل‌های بعدی با آن آشنا می‌شوید.
در توزیع‌هایی که چولگی زیاد نباشد رابطه تجربی زیر که به رابطه پیرسن معروف است برقرار می‌باشد.

$$\overline{X} - M \cong 3(\overline{X} - m)$$

مقدار عدم تقارن منحنی که با ضریب چاولگی سنجیده می‌شود را با نمودار نرمال مقایسه می‌کنیم برای بدست آوردن ضریب چاولگی از مفهوم گشتاور مرتبه kام طول میانگین (گشتاور مرکزی مرتبه kام) استفاده می‌کنیم. که عبارتست از :

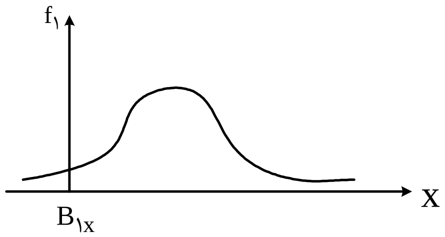
$$\mu_k = \frac{1}{N}\sum_{i=1}^{n} f_i \ (x_i - \overline{x})^k$$

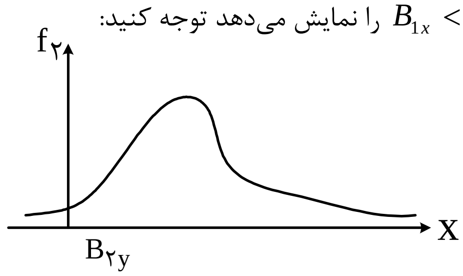
ضریب چاولگی که با B_1 نمایش داده می‌شود میزان عدم تقارن منحنی را نسبت به منحنی نرمال نمایش می‌دهد که به شکل زیر تعریف می‌شود:

$$B_1 = \frac{\mu_3}{\sqrt{\mu_2^3}} = \frac{\mu_3}{\delta^3}$$

در صورتی که منحنی متقارن باشد مقدار B_1 برابر صفر می‌باشد و هر چه B_1 مقدار بزرگتری داشته باشد نشان دهنده افزایش عدم تقارن می‌باشد.

به شکل زیر که دو مقدار $B_{1x} < B_{1y}$ را نمایش می‌دهد توجه کنید:





ضریب کشیدگی : نشان دهنده میزان کشیدگی منحنی می‌باشد و به صورت زیر تعریف می‌شود.

$$B_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\delta^4}$$

دو منحنی زیر متقارن می‌باشند اما ضریب کشیدگی در آنها متفاوت است.

$B_{2x} < B_{2y}$